

for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB227).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO322 [herein designated as UNQ283 (DNA48336-1309)] (SEQ ID NO:394) and the derived protein sequence for PRO322. It is understood that UNQ283 (DNA48336-1309) in fact encodes PRO322, and that SEQ ID NO:394 is a representation of the sequence based on sequencing techniques known in the art.

The entire nucleotide sequence of UNQ283 (DNA48336-1309) is shown in Figure 158 (SEQ ID NO:394). Clone UNQ283 (DNA48336-1309) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 166-168 and ending at the stop codon at nucleotide positions 946-948 (Figure 158). The predicted polypeptide precursor is 260 amino acids long (Figure 159). The full-length PRO322 protein shown in Figure 159 has an estimated molecular weight of about 28,028 daltons and a pI of about 7.87. Clone UNQ283 (DNA48336-1309) has been deposited with ATCC and is assigned ATCC deposit no. 209669.

Regarding the amino acid sequence of Figure 159, a potential N-glycosylation site is at amino acid 110 of SEQ ID NO:395. The serine proteases, trypsin family and histidine active site is identified at amino acids 69 through 74 of SEQ ID NO:395 and the consensus sequence is identified at amino acids 207 through 217 of SEQ ID NO:395. The kringle domain proteins motif is identified at amino acids 205 through 217 of SEQ ID NO:395. The putative signal peptide is encoded at about amino acids 1-23.

Analysis of the amino acid sequence of the full-length PRO322 polypeptide suggests that portions of it possess significant homology to neuropsin and other serine proteases, thereby indicating that PRO322 is a novel serine protease related to neuropsin.

EXAMPLE 63: Isolation of cDNA Clones Encoding Human PRO526

A consensus sequence was obtained relative to a variety of EST sequences as described in Example 1 above, wherein the consensus sequence obtained is herein designated DNA39626. Based on the DNA39626 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO526.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TGGCTGCCCTGCAGTACCTCTACC-3' (SEQ ID NO:401);

reverse PCR primer 5'-CCCTGCAGGTCATTGGCAGCTAGG-3' (SEQ ID NO:402).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA39626 consensus sequence which had the following nucleotide sequence:

hybridization probe

5'-AGGCACTGCCTGATGACACCTTCCGCGACCTGGGCAACCTCACAC-3' (SEQ ID NO:403).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO526 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal liver tissue (LIB228).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for

PRO526 [herein designated as UNQ330 (DNA44184-1319)] (SEQ ID NO:399) and the derived protein sequence for PRO526.

The entire nucleotide sequence of UNQ330 (DNA44184-1319) is shown in Figure 160 (SEQ ID NO:399). Clone UNQ330 (DNA44184-1319) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 514-516 and ending at the stop codon at nucleotide positions 1933-1935 (Figure 160). The predicted polypeptide precursor is 473 amino acids long (Figure 161). The full-length PRO526 protein shown in Figure 161 has an estimated molecular weight of about 50,708 daltons and a pI of about 9.28. Clone UNQ330 (DNA44184-1319) has been deposited with the ATCC on March 26, 1998. It is understood that the clone contains the actual sequence, whereas the sequences presented herein are representative based on current sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO526 polypeptide suggests that portions of it possess significant homology to the leucine repeat rich proteins including ALS, SLIT, carboxypeptidase and platelet glycoprotein V thereby indicating that PRO526 is a novel protein which is involved in protein-protein interactions.

Still analyzing SEQ ID NO:400, the signal peptide sequence is at about amino acids 1-26. A leucine zipper pattern is at about amino acids 135-156. A glycosaminoglycan attachment is at about amino acids 436-439. N-glycosylation sites are at about amino acids 82-85, 179-182, 237-240 and 423-426. A von Willebrand factor (VWF) type C domain(s) is found at about amino acids 411-425. The skilled artisan can understand which nucleotides correspond to these amino acids based on the sequences provided herein.

EXAMPLE 64: Isolation of cDNA Clones Encoding Human PRO531

An ECD database was searched and an expressed sequence tag (EST) from LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA was identified which showed homology to protocadherin 3. Based on this sequence, a search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequence. Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington).

A consensus DNA sequence was assembled relative to other EST sequences using phrap. Based on the consensus sequence obtained, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO531.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CTGAGAACGCGCCTGAAACTGTG-3' (SEQ ID NO:406);

reverse PCR primer 5'-AGCGTTGTCATTGACATCGGCG-3' (SEQ ID NO:407).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA sequence which had the following nucleotide sequence:

hybridization probe

5'-TTAGTTGCTCCATTGAGGAGGATCTACCCTTCCTCTGAAATCCGCGGAA-3' (SEQ ID NO:408).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO531 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal brain tissue (LIB153). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO531 [herein designated as UNQ332 (DNA48314-1320)] (SEQ ID NO:404) and the derived protein sequence for PRO531.

The entire representative nucleotide sequence of UNQ332 (DNA48314-1320) is shown in Figure 162 (SEQ ID NO:404). It is understood that the actual sequence is that within the clone deposited with the ATCC as DNA48314-1320. Clone UNQ332 (DNA48314-1320) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 171-173 and ending at the stop codon at nucleotide positions 2565-2567 (Figure 162). The predicted polypeptide precursor is 789 amino acids long (Figure 163). The full-length PRO531 protein shown in Figure 163 has an estimated molecular weight of about 87,552 daltons and a pI of about 4.84. Clone UNQ332 (DNA48314-1320) has been deposited with the ATCC on March 26, 1998.

Analysis of the amino acid sequence of the full-length PRO531 polypeptide suggests that portions of it possess significant homology to protocadherin 3. Moreover, PRO531 is found in the brain, like other protocadherins, thereby indicating that PRO531 is a novel member of the cadherin superfamily.

Still analyzing the amino acid sequence of SEQ ID NO:405, the cadherin extracellular repeated domain signature is found at about amino acids 122-132, 231-241, 336-346, 439-449 and 549-559 of SEQ ID NO:405. An ATP/GTP-binding site motif A (P-loop) is found at about amino acids 285-292 of SEQ ID NO:405. N-glycosylation sites are found at least at about amino acids 567-570, 786-790, 418-421 and 336-339 of SEQ ID NO:405. The signal peptide is at about amino acids 1-26, and the transmembrane domain is at about amino acids 685-712 of SEQ ID NO:405.

EXAMPLE 65: Isolation of cDNA Clones Encoding Human PRO534

A consensus sequence was obtained relative to a variety of EST sequences as described in Example 1 above, wherein the consensus sequence obtained is herein designated DNA43038. Based on the 43048 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO534.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CACAGAGCCAGAAGTGGCGGAATC-3' (SEQ ID NO:411);

reverse PCR primer 5'-CCACATGTTCTGCTCTTGTCTCTGG-3' (SEQ ID NO:412).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA43038